Learning and Individual Differences 59 (2017) 127–140

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Predicting the integrated development of word reading and spelling in the early primary grades☆

Moniek M.H. Schaars*, Eliane Segers, Ludo Verhoeven

Behavioural Science Institute, Radboud University, The Netherlands

ARTICLE INFO

Keywords:
Spelling development
Word reading development
Precursors
Integrated relationship
Structural equation modeling

ABSTRACT

Word reading and spelling processes are assumed to be highly related to each other and to early literacy measures. However, the debate on how reading and spelling interact in early development is far from resolved yet. The present study examined the singular and integrated word reading and spelling development during the first two grades of primary education in relation to kindergarten precursor measures of short-term memory, vocabulary, rapid naming, and early literacy (phonemic awareness, grapheme-to-phoneme knowledge) in 487 Dutch children. Structural equation models showed that word reading and spelling development separately were highly stable and consistently autoregressive in nature during first and second grade. Both word reading and spelling development were predicted by early literacy, and word reading development was additionally predicted by rapid naming. An integrated model for word reading and spelling development showed that word reading skill predicted subsequent spelling skills in Grade 2 over and above the autoregressive prediction. No reciprocal relation of spelling to subsequent word reading has been found.

1. Introduction

One major job for children in elementary school is the development of proper literacy skills. Reading and spelling are two core components of literacy. It has been suggested that reading and spelling derive from the same cognitive and linguistic processes (e.g., Caravolas et al., 2012; Juel, Griffith, & Gough, 1986; Landerl & Wimmer, 2008; Shanahan, 1984). Theoretical models concur with the idea that orthographic, phonological, and semantic components are involved in both reading and spelling processes (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Frith, 1985; Plaut, McClelland, Seidenberg, & Patterson, 1996; Van Orden, Pennington, & Stone, 1990). Indeed, behavioral studies have shown that reading and spelling are highly related (e.g., Ehri, 1989; Juel, 1988), and neuroimaging studies have provided evidence that reading and spelling activate overlapping brain regions (Pugh et al., 2006). Despite this strong suggestion of relatedness, there are only a few longitudinal studies about how the developmental pathways of reading and spelling are related in the early elementary grades. Also, how reading and spelling can be predicted from kindergarten precursor measures of phonemic awareness, grapheme-to-phoneme knowledge, rapid naming, vocabulary and short-term memory needs further investigation. It is noteworthy that only a few studies have combined reading and spelling development as well as their precursors in one integrated model. Furthermore, such studies have hardly been conducted in relatively transparent orthographies in which reading fluency is a better measure than reading accuracy to establish reading ability. Although it seems a matter of course that reading and spelling are somehow related, the underlying nature of this relation has not yet been clarified. Therefore, the present study aimed to describe the early singular and integrated word reading efficiency and spelling development in the first two primary grades in relation to kindergarten precursors in the relatively transparent Dutch orthography. This large longitudinal Dutch study contributes to the knowledge about the general underlying principles in literacy development.

1.1. Word reading development and its precursors

Word reading development has generally been described as a phase-like model (Ehri, 2005; Frith, 1985). During a first, phonologically driven, decoding phase, children explicitly learn to accurately decode written words into their auditory counterparts by the one-to-one conversion of graphemes into phonemes (Coltheart et al., 2001). After acquiring these elementary decoding skills, children gradually learn to read more complex and longer words containing orthographic structures, for example, consonant clusters and multi-syllables. Every time children encounter a specific internal structure, this larger unit becomes

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author at: Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands.
E-mail address: m.schaars@pwo.ru.nl (M.M.H. Schaars).

http://dx.doi.org/10.1016/j.lindif.2017.09.006
Received 23 December 2016; Received in revised form 31 August 2017; Accepted 10 September 2017
1041-6080/

better consolidated in an internal orthographic lexicon. Consequently, the orthographic lexicon becomes better specified (Perfetti, 1992). By this self-teaching mechanism, beginning readers gradually become more efficient and fluent (Share, 1999; Tucker, Castles, Laroche, & Deacon, 2016), and the connections between the orthographic (graphemes), phonological (phonemes), and semantic (word meanings) components become stronger, as proposed in the Phonological Coherence model (Bosman & Van Orden, 1997; Van Orden et al., 1990).

High individual stability over time has been evidenced for word reading development in both transparent (e.g., Schaars, Segers, & Verhoeven, 2017; Verhoeven & Van Leeuwe, 2009) and more opaque orthographies (e.g., Caravolas, Lervåg, Defior, Málková, & Hulme, 2013; Furnes & Samuelsson, 2011; Juul, Poulsen, & Elbro, 2014; Steacy, Kirby, Parrila, & Compton, 2014) and precursors of word reading development are well established. Phonemic awareness, grapheme-to-phoneme correspondences, rapid naming, and vocabulary have been found to be relevant precursors of word reading development (e.g., Al Otaiba & Fuchs, 2002; Kirby, Georgiou, Martinussen, & Parrila, 2010; Melby-Lervåg, Lyster, & Hulme, 2012; Moll et al., 2014; Nelson, Benner, & Gonzalez, 2003). Recently, also individual variation in visual and verbal short term memory have been shown to contribute to the prediction of later reading performances (Bosse & Valdois, 2009; Van den Boer, De Jong, & Haentjens-van Meeteren, 2013). The relative contribution of precursors might differ between developmental phases and orthographies, with rapid naming as an especially important predictor of reading efficiency in transparent orthographies (Babayiğit & Stainthorp, 2010; Caravolas et al., 2013; De Jong & Van der Leij, 1999; Ziegler & Goswami, 2005, 2006).

1.2. Spelling development and its precursors

Regarding the development of spelling, a few longitudinal studies have been conducted. Most studies, especially on specific difficulties in the spelling system, have been conducted in the English orthography (e.g., Treiman, Cassar, & Zukowski, 1994; but see Caravolas, 2004). However, also in transparent orthographies an autoregressive developmental spelling path has been evidenced, meaning that the individual differences of spelling ability seem to be largely preserved over time. A Norwegian longitudinal study of Lervåg and Hulme (2010) showed for example that, although children varied in how fast they learned to spell words, these individual differences could best be described as variations around a single trajectory. A Dutch cross-sectional study (second to sixth grade) of Keuning and Verhoeven (2008) also showed that spelling development can be best described in terms of a stable continuous learning process. Although literature agrees on a certain autoregressive development of spelling skills, the autoregression is assumed to be less consistent as compared to reading development (Desimoni, Scalisi, & Orsolini, 2012; Pinto, Bigozzi, Tarchi, Gamannossi, & Canneti, 2015).

With regard to the precursors of spelling, converging evidence indicates that as in word reading, phonemic awareness and knowledge of grapheme-to-phoneme correspondences are at least as important (e.g., Caravolas, Hulme, & Snowling, 2001; Furnes & Samuelsson, 2010; Nikolopoulos, Goulandris, Hulme, & Snowling, 2006; Torppa, Georgiou, Niemi, Lerkkanen, & Poikkeus, 2016). In addition, studies on children with dyslexia showed that cognitive and linguistic skills that are important in reading, are also contributing in spelling skills (e.g., Morken & Helland, 2013). However, reading and spelling development are, at least partially, based on different compositions of cognitive and linguistic determinants (Ahmed, Wagner, & Lopez, 2014; Babayiğit & Stainthorp, 2010; Caravolas et al., 2001; Caravolas et al., 2012; Nikolopoulos et al., 2006; Torppa et al., 2016). In addition, different compositions have been shown between different orthographies (e.g., Furnes & Samuelsson, 2009; Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012). Vaessen and Blomert (2013) found that phonemic awareness and grapheme-phoneme knowledge were stable predictors of spelling in Dutch, whereas their contribution to reading decreased during development. Vaessen and Blomert used a cross-sectional study design in which only concurrent relations between predictors and reading fluency were studied, making interpretations about causality to be taken with caution. Their results do add to the suggestion that the connection between phonology and orthography (see Bosman & Van Orden, 1997; Van Orden et al., 1990) remains more important for spelling than for word reading during development. Babayiğit and Stainthorp (2010, in Turkish) also showed higher predictive power of phonological awareness to spelling skills as related to word reading skills.

Vaessen and Blomert (2013) found no contribution of rapid naming to spelling development, whereas the contribution of rapid naming to reading was relatively strong. Although rapid naming has previously been found to be a predictor of spelling ability (Caravolas et al., 2012; Furnes & Samuelsson, 2010; Verhagen, Aarnoutse, & Van Leeuwe, 2010), it has been proposed to be more related to reading skills, since fluent reading is a timed performance from the very beginning (at least in a transparent orthography) whereas spelling is not (e.g., Kirby, Desrochers, Roth, & Lai, 2008; Lervåg & Hulme, 2010). Also, the contribution of short term memory (e.g., Lervåg & Hulme, 2010) and vocabulary (e.g., Verhagen et al., 2010) have previously been evidenced in the prediction of spelling abilities. The Norwegian study of Lervåg and Hulme (2010) longitudinally examined all the before mentioned cognitive and linguistic contributions to spelling development in one and the same study. They found that grapheme-to-phoneme knowledge and phonemic awareness (which could hardly be differentiated from each other) consistently were the most powerful predictors of spelling development. Other studies agree on the contribution of phonemic awareness and grapheme-to-phoneme knowledge to spelling performances (e.g., Caravolas et al., 2001; Hulme, Snowling, Caravolas, & Carroll, 2005; Muter, 1998) and it is in line with the Phonological Coherence model (Bosman & Van Orden, 1997).

1.3. The integrated development of word reading and spelling and their precursors

Word reading and spelling skills have long been considered more or less the same skills, performed in opposite directions (Ehri, 2000; Perfetti, 1997). Similar fundamental skills would be underlying to the performance of both word reading and spelling, in that view. More recently, however, it has been argued that spelling is not a one-to-one reversal of word reading, although word reading and spelling both rely on knowledge of the alphabetic principle (Abbott, Berninger, & Fayol, 2010; Foorman, Arndt, & Crawford, 2011; Shanahan, 2006). The Phonological Coherence model of Bosman and Van Orden (1997) shows a network with recurrent relations between phonemic, graphemic and semantic information. All relations can be activated in both directions, meaning that both conversion from graphemes-to-phonemes and phonemes-to-graphemes are supported in this model.

Spelling, however, is argued to be more difficult than reading (Bosman & Van Orden, 1997). One reason is because inconsistencies in spelling must be resolved with weaker cues of grapheme-semantic relations, whereas inconsistencies in reading can rely on stronger phoneme-semantic cues (Bosman & Van Orden, 1997). In other words, correct spelling requires the active generation of an orthographic structure, whereas reading basically requires its identification and recognition (Fletcher-Flinn, Shankweile, & Frost, 2004). A second reason is because, in general, there are more graphemes to choose from for writing down a phoneme, than there are phonemes for pronunciation of a grapheme. As a consequence of this asymmetry between the regularity of phoneme-to-grapheme conversion as compared to grapheme-to-phoneme conversion, individual differences in children's spelling skills are larger than those in reading skills from the very beginning. A brain study in poor German spellers with normal reading skills (Gebauer et al., 2012) showed that the grapheme-to-phoneme knowledge of the children could compensate for weak orthographic representations in reading. However, due to asymmetries in conversion, this could not compensate in spelling. Other German studies on poor spellers found similar results (e.g., Moll & Landerl, 2009; Wimmer & Mayringer, 2002). Although orthographies differ in the reciprocity of the conversion (e.g., Finnish and Turkish show high regularity in both directions; Babayiğit & Stainthorp, 2010; Leppänen, Nieme, Aunola, & Nurmi, 2006), a certain asymmetric consistency is typical across alphabetic orthographies (see Van Orden, Bosman, Goldinger, & Farrar, 1997, and e.g., Bekebrede, Van der Leij, & Share, 2009 (Dutch); Pinto et al., 2015 (Italian); Stone, Vanhoy, & Van Orden, 1997 (English); Wimmer & Mayringer, 2002 (German); Ziegler, Jacobs, & Stone, 1996 (French)). Finally, writing down a heard word requires detailed phonological analysis of speech, which is of less prominent relevance in reading a written word (Carroll, Snowling, Stevenson, & Hulme, 2003).

Although a certain relationship between word reading and spelling is assumed, the aforementioned differences point to an asymmetry in the skills and their relative development. A prominent question is how both skills might contribute to each other's development. The exact nature of the relationship between reading and spelling development has, however, been investigated in only few studies and no consensus has been reached so far. Studies showed that reading and spelling are also influencing each other's development. Both bidirectional (prediction of reading-to-spelling and spelling-to-reading; e.g., Abbott et al., 2010) and unidirectional (e.g., one-way reading-to-spelling prediction; e.g., Ahmed et al., 2014) influences of word reading and spelling have been assumed. Ellis and Cataldo (1990) conducted a longitudinal study in an English Grade 1 sample to explore the direction of the relationship. Their path model showed a bidirectional influence, with important influence of spelling on early reading and only a weak influence of reading on early spelling. A cross-sectional study of Abbott et al. (2010) also found reciprocal influence from reading-to-spelling and from spelling-to-reading during second to seventh grade in the US, albeit they initially found an unidirectional influence of spelling on word reading from first to second grade. Although the influences of reading to spelling and spelling to reading were significant, they were rather small (coefficients ranging from 0.14–0.33; $M = 0.22$) in comparison to the autoregression of the skills (coefficients ranging from 0.56–0.83; $M = 0.68$). The longitudinal study of Ahmed et al. (2014), which was conducted from first to fourth grade in the US, showed best fit for models with reading-to-spelling influences as compared to models with spelling-to-reading influences or bidirectional models. In line with this best fit, the study of Caravolas et al. (2001) showed a considerable unidirectional influence of emergent reading accuracy to emergent spelling (standardized coefficient of 0.46–0.47, over and above the autoregression of spelling being 0.36), while no evidence for the reversed influence of spelling on reading development was found. Their study was conducted in the UK from halfway of the second kindergarten year to halfway Grade 2, so floor levels were scored on both reading and spelling during the first and second measurement moment.

The studies on the integrated literacy development described above have all been conducted in the outlier English orthography. Few studies in more transparent orthographies have been done to disentangle what of these findings can be considered general underlying principles in literacy development. Lerkkanen, Rasku-Puttonen, Aunola, and Nurmi (2004) collected a longitudinal dataset in 83 first grade children in the transparent Finnish orthography. In line with the English studies, their results showed that reading and spelling both showed substantial stability. Reading and spelling were bidirectionally related in the first months of first grade (standard coefficients of reading to spelling varying between 0.27 and 0.31; spelling to reading 0.23–0.51). However, only the reading-to-spelling prediction remained during the second half year of first grade. In a follow up cross-lagged longitudinal study, Leppänen et al. (2006) studied the integrated development of reading and spelling from kindergarten to becoming literate (start of Grade 2). The model of Leppänen et al. also showed high stability for both reading and spelling, and they found a contribution of preschool spelling abilities to reading during the very beginning of becoming literate (with standardized coefficients of 0.21 and 0.38). They explained this association by stating that the emerging preschool spelling skills are highly comparable with phonological awareness skills (also suggested by Babayiğit & Stainthorp, 2010), which are assumed to be predictive for reading development. Soon, only the reading-to-spelling contributions remained in their longitudinal model in Grade 1 and the start of Grade 2. Both Lerkkanen et al. and Leppänen et al. covered a broader field of reading development, since they also included reading comprehension in their latent reading measure. They did, however, not take into account reading speed or efficiency in their study. In line with the eventual unidirectional finding in the Finnish studies, in an accuracy focused study among 170 Italian children (transparent orthography; Desimoni et al., 2012), it has also been found that reading errors predicted spelling errors in a unidirectional relationship. Desimoni et al. interpreted theses outcome by stating that having correct phonological reading skills enhances correct spellings of transparent words. The study analyzed the relation between two time points from Grade 1 to Grade 3. This covers a large measurement interval without intermediate measurement moments from the phase during which both reading and spelling rely heavily on correct conversion of graphemes and phonemes towards a more orthographically driven literacy phase (both in reading and spelling). Therefore, it is not clear how interactions between word reading and spelling process during this phase. Another Italian study (Pinto et al., 2015) found bidirectional influences of reading (both accuracy and speed) and spelling in a free writing task. They studied relations between two time points during Grade 1 and Grade 2 for 57 children.

Although all mentioned studies did find an influence of reading-to-spelling, the additional presence of the reversed influence of spelling-to-reading is not consistently found. Furthermore, no consensus has been reached yet about the consistency and the strength of the influences. Differences in the directionality of the findings can be partly explained by characteristics of the studied orthographies. Many so-called transparent orthographies show asymmetries in the regularity of grapheme-to-phoneme conversion and phoneme-to-grapheme conversion. Less regularity in the phoneme-to-grapheme conversion direction as compared to the grapheme-to-phoneme conversion, stimulates the use of orthographic strategies in spelling already during early development. Therefore, it is suggested that in relatively transparent (but asymmetric) orthographies, the interaction between reading and spelling development is different from the integrated development in other, generally opaque, orthographies.

Differences in the directionality that is found in previous literature is not only due to differences in orthographic systems, however. In addition, differences come from variations in measurement intervals and periods of interest, and differences in constructs that have been measured. Frith (1985) was one of the first to hypothesize that period of interest might influence findings, since she suggested that phase-wise development of reading and spelling might not run in synchrony. In her theoretical model, reading and spelling shift their leading role in different developmental phases, and therefore, influence each other's development in different developmental phases. For example, children might first practice their alphabetic strategies in spelling before they start to apply this practiced skill to reading too. The orthographic strategy, in contrast, might develop first in reading before children adopt the strategy in spelling too. This implies that reading development serves spelling development as soon as additional orthographic strategies come into play. Bosman and Van Orden (1997) elaborated on that theory by adding that the differences and asymmetries between reading and spelling development are enhanced by how we use reading and spelling skills in daily life. Usually, people read more than they spell. It has indeed been argued (e.g., Cunningham & Stanovich, 1990; see also Ellis, 1994) that fluent readers read more, resulting in better specified orthographic lexicons, which on its turn can be used in spelling, even in young children during early development.

Regarding precursors in the integrated literacy development, only few studies simultaneously studied the influence of cognitive and linguistic precursors in an integrated model of both spelling and reading. Caravolas et al. (2001) found that, after the autoregressive and interdependent relations were taken into account, grapheme-to-phoneme knowledge and phoneme isolation skill were the only unique predictors of spelling, whereas reading was uniquely predicted by letter name knowledge and phoneme isolation skill. It should, however, be taken into account that in the Caravolas' study, floor levels were scored on the first two measurement moments of spelling. This might have influenced their results. Leppänen et al. found that phonological awareness and grapheme-to-phoneme knowledge measured in preschool years uniquely impacted both the reading accuracy (including reading comprehension) and spelling development during preschool. Grapheme-to-phoneme knowledge even had an additional impact on reading development in Grade 1. Both studies have not included precursor measures of short term memory and vocabulary. In addition, reading fluency and lexical retrieval, like rapid naming, was not taken into account in previous integrated models. Rapid naming has previously been established as a relevant precursor of reading development (e.g., Moll et al., 2014). In fact, rapid naming was currently proposed to represent orthographic processing and, therefore, to influence the building of an orthographic lexicon (see Moll et al., 2014 for an overview of the literature). It has been proposed, therefore, to be related to spelling processes as well.

It is far from clear how the integrated development in early reading and spelling unfolds in transparent orthographies. The current study investigated how early reading efficiency and spelling interact with each other in Dutch children, and how this integrated development is fed by cognitive and linguistic underpinnings, including rapid naming. Although the few studies on the integrated development of reading and spelling all mention interdependencies, mostly with a predicting power in the reading-to-spelling direction, no consistency about the exact nature of the interrelationship has been reached. First, it remains far from clear how the word reading and spelling development continues from the literacy development in Grade 1 to becoming more proficient in Grade 2. Second, truly longitudinal consideration in one and the same cross-lagged design is rare for both the early word reading and spelling development. Third, most previous research on the relation between reading and spelling, even the study in the transparent Finnish orthography, have measured reading accuracy instead of reading efficiency (but see e.g., Pinto et al., 2015). In an efficiency measure, the accuracy is inherently taken into account, but the focus is on fluency. In English studies, early word reading performance is best measured in terms of accuracy, because children simply cannot yet efficiently read during the initial phases of word reading development. The reading accuracy data in the Finnish study (Lerkkanen et al., 2004) were highly skewed, for example, showing that accuracy as an indicator is less useful for discriminating poor readers from good readers in transparent orthographies (Babayiğit & Stainthorp, 2010; Landerl & Wimmer, 2008; Pinto et al., 2015; Verhoeven & Van Leeuwe, 2009). Lastly, only few studies have taken cognitive and linguistic precursors into account in studying the nature of the integration of word reading and spelling development.

1.4. Present study

In the present study, we examined the singular and integrated development of both word reading and spelling in Dutch during the first two years of formal instruction in terms of their cognitive and linguistic determinants. Therefore, the current study adds to our knowledge on the integrated early reading and spelling development in different orthographies, strengthening our knowledge about general underlying principles in literacy development. This longitudinal study was conducted from kindergarten up to the end of second grade, containing 487 Dutch participants. Cognitive and linguistic measures were assessed in kindergarten, and word reading and spelling performances were assessed four times during first and second grade (halfway and by the end of the school years). We aimed to answer to the following questions:

1. How is the singular development of early word reading and of spelling in Dutch determined by cognitive and linguistic kindergarten measures, and is this different for reading and spelling?
2. How are early word reading and spelling development in Dutch related to each other and how is the integrated model determined by cognitive and linguistic kindergarten measures?

We expected that both word reading and spelling development could be predicted from the kindergarten cognitive and linguistic measures phonemic awareness, grapheme-phoneme knowledge, and rapid naming, and, to a lesser extent, by short term memory and vocabulary. We expected that both word reading and spelling development would be highly autoregressive in nature. We further expected that word reading and spelling development would be interrelated to each other, with a higher influence of reading-to-spelling as compared to spelling-to-reading. Both reading and spelling experience contributes to the consolidation of orthographic representations in the mental lexicon. This can be used in the accurate and automated retrieval during both reading and spelling performances. However, the connections of graphemes-to-phonemes are assumed to be stronger and more consistent as compared to the phoneme-to-grapheme connections (Van Orden et al., 1990), adding more to the development of stable orthographic representations. Less influence of spelling-to-reading was expected, since actual word decoding would take strong account for its own further development (and the consolidation of the orthographic lexicon). Furthermore, in the Dutch orthography it could be hypothesized that learning to read is in the forefront of learning to spell. Therefore, early benefits from one skill to the other, flows more logically from word reading to spelling. See Fig. 1 for a conceptual model showing the relations to be tested in the current integrated study on early literacy development.

2. Method

2.1. Participants

The current study is part of a larger longitudinal cohort study on literacy development. A sample of 37 general education primary schools throughout the Netherlands participated in the larger longitudinal study. Our sample was treated in accordance with institutional guidelines as well as with APA ethical standards. Schools, parents, and children were informed about the purpose of the research, the expected duration of the experiments, and the procedures. They were informed about whom to contact for questions about the research. Schools gave active consent to participate in the longitudinal study. Both schools (as institution) and parents (of individual participants) were aware of their right to decline participation and to withdraw from the research any time before or during the research project. After each academic year, the schools were asked if they were willing to maintain their participation. The data collection focused on normal educational practices, curricula, and methods in daily educational settings. Schools were debriefed with information about the results and conclusions of the research. The children were recruited in the end of kindergarten and, for the current study, further monitored until the end of Grade 2. We obtained passive informed consent from the parents of 1006 children. Children who were expected to stay in kindergarten for an extra year were excluded beforehand from participation in the cohort study.

All schools in this study made use of the same highly systematic

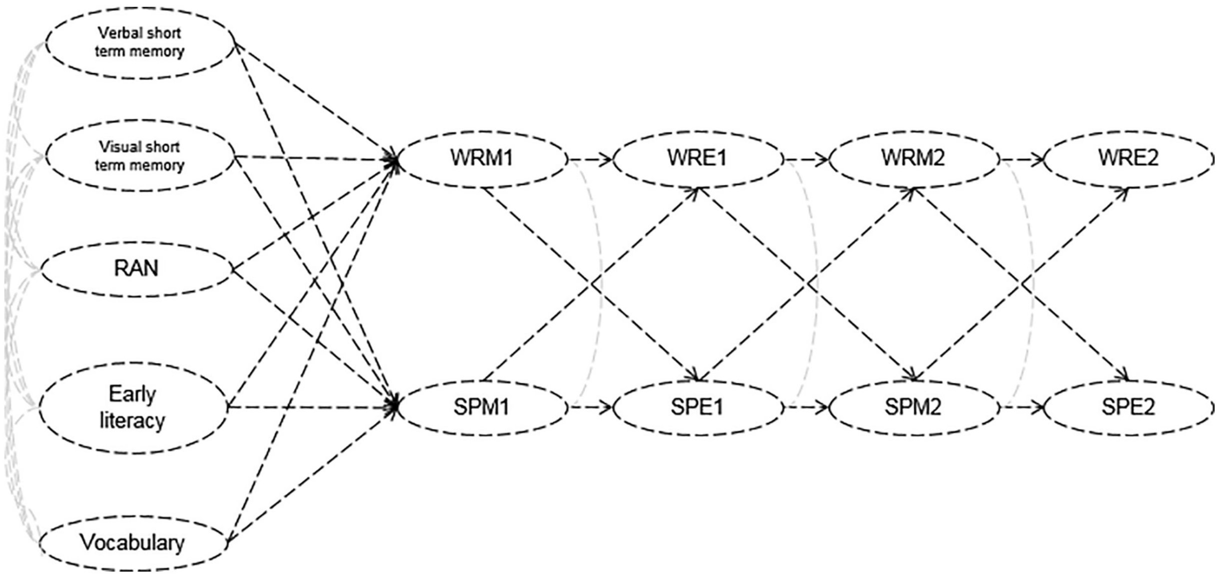

Fig. 1. Conceptual model of tested relations in the integrated literacy development in Dutch.

phonics based reading method in Grade 1 (*Veilig leren lezen*; '*Learning to read safely*'; Mommers et al., 2003), by which over 80% of the children in the Netherlands learn to read.

In total, 15 schools were excluded for analyses in the current study, because they made the transition to another version of the standardized spelling measurement procedure somewhere during Grade 1 or Grade 2. The 22 schools that continued to use the same standardized spelling measurement procedure (De Wijs, Kamphuis, Kleintjes, & Tomesen, 2010) were included in the current sample. The schools varied in size and both rural and urban areas were represented. All regions of the Netherlands were represented in the sample.

For current analyses, children who missed two or more of the repeated standardized measures of word reading and/or spelling were excluded from the initial sample (74 children; 13% of the participants). The exclusions were mostly due to movements or transfer to other schools, or to long absence for illness during the longitudinal study. Analyses were conducted with a representative sample of 487 Dutch children (246 boys; 241 girls). In the Netherlands, kindergarten is a two-year program prior to first grade. The children were firstly assessed by the end of kindergarten. The mean age of the children at that moment was 6;2 years (SD = 0;4). All children spoke Dutch and 18% spoke another language at home as well. This percentage is representative to the multicultural nature in the Netherlands. Different social classes were represented by the sample, as indicated by the educational level of the main care giver. The distribution was representative for the distribution of educational level in the Netherlands (Centraal Bureau voor de Statistiek [Statistics Netherlands], 2013).

2.2. Materials

2.2.1. Cognitive and linguistic skills in kindergarten

The kindergarten test battery on child characteristics consisted of seven tasks, which were designed for the purpose of the longitudinal study on literacy development (Schaars et al., 2017). The tasks measured different constructs that were assumed to be involved in learning to read and spell. All tasks included items increasing in difficulty, except for the tasks measuring rapid naming and grapheme-phoneme knowledge. Difficulty was established based on length and CV-structure (Schreuder & Van Bon, 1989), phoneme position (De Graaff, Hasselman, Verhoeven, & Bosman, 2011), and phoneme characteristics and sound assimilation effects (Geudens & Sandra, 2003). Two practice items preceded each task, except from the grapheme-phoneme knowledge task. All pictures in the rapid naming task were practised. The score on each task was the number of correct responses. To get insight into the psychometrical quality of the kindergarten test, the Cronbach's alpha of each task was normed on the 1006 children that initially participated in the kindergarten test wave. In addition, the variance from the mean and the deviation of the scores between student's provided good distributions for sensible analyses of the individual variations among the children. No floor or ceiling levels for the tasks were reached for children at the end of kindergarten.

2.2.1.1. Initial phoneme isolation. In order to examine initial phoneme isolation skills, the child had to sound out the first phoneme of 10 orally introduced monosyllabic CVC-structured words (e.g., *m̲uis*, *s̲oep*). The task had good reliability (Cronbach's $\alpha = 0.83$).

2.2.1.2. Word segmentation task. To assess word segmentation skills, the child had to serially pronounce each phoneme of an orally introduced word. The task contained 10 words with increasing difficulty, starting with CVC-structured words followed by CCV(C)- or (C)VCC-structured words and CCCV(C)- or (C)VCCC-structured words. The task was discontinued after five consecutive incorrect responses. It was assumed that further items could be considered incorrect. The task had good reliability (Cronbach's $\alpha = 0.85$).

2.2.1.3. Grapheme-phoneme knowledge. To examine grapheme-to-phoneme conversion skill, the child was asked to sound out 34 graphemes used in Dutch. The graphemes c, q, x, and y were excluded from this task, because these graphemes are very low frequent in the Dutch reading system. The graphemes were introduced in columns on a card. The font corresponded with the font used in the Grade 1 reading curriculum, and graphemes were introduced in lower case. In this task only the grapheme sound was considered correct; naming the grapheme's name was incorrect. The task had excellent reliability (Cronbach's $\alpha = 0.93$).

Since we expected a high interrelationship between the three measures of early literacy (Caravolas et al., 2001; Hulme et al., 2005; Lervåg & Hulme, 2010; Muter, 1998), we conducted a principal axis factor analysis with oblique rotation (Promax; Muthén & Muthén, 2007) on the three measures described above. The analysis revealed one component with high loadings (0.81 to 0.83) that explained 66.96% of the variance. The Kaiser-Meyer-Olkin measure verified the adequacy of this analysis, KMO = 0.69 (middling; Hutcheson & Sofroniou, 1999).

All analyses in the current study were conducted using a latent factor variable named Early literacy.

2.2.1.4. Rapid naming of objects. The Rapid naming task was developed to measure the lexical retrieval speed of visually presented objects. Nonalphanumeric stimuli are preferred in prereaders, since the stimuli should be 'highly familiar' to tap into the automated retrieval skills (e.g., Kirby et al., 2010; Lervåg & Hulme, 2009). Furthermore, Landerl et al. (2013) found in their regression models that both digits and pictures were reliable predictors of diagnostic status. The task consisted of five different pictures, all corresponding with one-syllable high frequent Dutch words (viz., *saw*, *pot*, *thumb*, *trousers*, *tent*) which were familiar to children in the age of 6 (Schaerlaekens et al., 1999). The pictures were presented in six columns of 22 objects (132 objects in total). The task was preceded by a short practice session to make sure the child named the presented pictures correctly. The child had to name as many objects as possible from top to bottom during 1 min. The task had excellent reliability (Cronbach's $\alpha = 0.95$).

2.2.1.5. Verbal short term memory. The quality of the phonological store of verbal short term memory was measured with a pseudoword repetition task. There were 20 one-to-four syllable pseudowords in this task, introduced by the test assistant in ascending order of length. The child was asked to accurately repeat each word. The whole word had to be repeated correctly; stress differences and substitutions due to certain articulation errors in individuals were counted as correct. After five consecutive incorrect responses, the task was discontinued. The reliability of this task was good (Cronbach's $\alpha = 0.77$).

2.2.1.6. Visual short term memory. The task on visual short term memory measured the sequential short term memory of concrete visual information. The child was asked to remember the order of series of visually presented figures (viz., *fish*, *cow*, *ship*, *chicken*, *sock*) that were presented for 5 s by the test assistant. After those 5 s, the booklet with the pictures closed and the child was asked to put cards with the pictures in the same order as had been presented in the booklet. The task contained 15 series. The number of figures in a series increased from two to five figures to remember. After three consecutive incorrect series, the task was discontinued. The reliability of this task was good (Cronbach's $\alpha = 0.77$).

2.2.1.7. Vocabulary. The vocabulary task was developed to measure the active vocabulary of the children. The pictures in the task were extracted from a Dutch language test for foreign children (TAK; Verhoeven & Vermeer, 1986). The task contained both nouns and verbs. Twenty-nine black and white line pictures were shown to the child. Every picture was accompanied by a little phrase pronounced by the test assistant. The child was asked to complete the phrase by naming the correct word. The task was discontinued after five consecutive incorrect responses. The task had good reliability (Cronbach's $\alpha = 0.83$).

2.2.2. Word reading development

2.2.2.1. Standardized word reading measures. With standardized tests, we systematically assessed children's ability to decode words (*Drie-minutentoets*; "Three-minutetest"; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). The total task consisted of three word cards which varied systematically with regard to orthographic transparency (cf. Nunn, 1998): the first card consisted of transparent one-syllable simple-structured words; the second card contained transparent words with at least one consonant cluster; the third card contained highly transparent words with at least two syllables. Per card, the child was asked to accurately read as many words as possible during 1 min. The amount of correct read words was the score on a reading card. The sum score of both cards was the word reading efficiency score. This combination of card scores was considered reliable, with a Cronbach's α of 0.96 for the sum of the two cards halfway Grade 1 and 0.97 for the combination of three cards in subsequent measurement moments (Krom et al., 2010).

2.2.3. Spelling development

2.2.3.1. Standardized spelling measures. We administered spelling ability with standardized tasks that vary and increase in difficulty over the measurement moments (*CITO Spelling toets*; "CITO Spelling assessments"; De Wijs et al., 2010). The standardized spelling tasks were based on the Item Response model of Rasch (1960). Scores on the subsequent test moments were converted to standardized scores, following the principles of the One Parameter Logistic Model (Verhelst & Glas, 1995; Verhelst, Glas, & Verstralen, 1995). This made children's performances comparable to each other (between individuals), and to previous performances (within individuals). Administration of the same spelling test on successive measurement occasions was clearly not an option because of learning and memory effects on the specific items. In addition, it is not meaningful to administer the same spelling test in the lower as well as the upper grades of elementary school. Standardized scores were calculated based on national norm scores of a representative group of Dutch children (ascending from national average score of 106 halfway Grade 1 up to 120 at the end of Grade 2; De Wijs et al., 2010). The norms differed for each subsequent measurement moment, meaning for example that a higher standardized score was expected for children in the end of Grade 2 as compared to the scores halfway Grade 2. Three different spelling tasks were used in the current study.

2.2.3.1.1. Single word dictation task. In the single word dictation task, 25 single words were pronounced. In an examination booklet, a little black line picture accompanied each orally pronounced target word. The children were asked to write down the pronounced words next to the line pictures. The amount of correct spelled items is the score on the single word dictation task.

2.2.3.1.2. Sentence context dictation task. In this task, 25 target words were each orally represented in a short sentence and the children were asked to write down the target word using paper and pencil. An example of one item is: "The dog is playing with the ball"… write down… "ball". The amount of correct spelled items is the score on the sentence dictation task.

2.2.3.1.3. Spelling decision task. In the decision task, four written sentences are represented per item. The four sentences all contain one word in bold. One of the four bold words is spelled inaccurately. The child was asked to encircle the word with the inaccurate spelling. In total, 25 items were assessed. The amount of correct encircled inaccurate spellings is the score on the decision task.

The single word dictation, the sentence context dictation, and the spelling decision task are all indicators of the same underlying ability (De Wijs et al., 2010), and the standardized tasks were all reliable, with all Cronbach's $\alpha > 0.90$ (De Wijs et al., 2010).

2.3. Procedure

At the end of Kindergarten, children participated in an individual assessment of about 30 min. The tasks in the test battery were conducted by the first author and eight trained test assistants with a Bachelor's or Master's degree in educational, psychological or linguistic sciences. Tasks were conducted in the same fixed order for all children. The test assistant orally provided instructions for all tasks. Other than for the practice items, no feedback on the correctness of item scores was provided to the children. All tasks were administered individually in a quiet room at school during regular school hours.

Word reading efficiency, taking into account both accuracy and speed, was administered in a standardized task halfway and by the end of both Grade 1 and Grade 2 (in February and in June). In total, four repeated measures of word reading ability were administered during the two school years. Only the first two cards of the standardized word reading task were administered halfway Grade 1, because the children were not yet able to read the words on the third card. In all subsequent measurement moments, all three cards were administered. This procedure is in accordance with the manual of the standardized task (Krom et al., 2010). The tests were carried out individually by certified teachers of the participating schools (mostly the remedial teacher of the schools) in a quiet room within the school building. The children received a short instruction before the test started.

In accordance with the standardized assessment of word reading, the children were assessed on two occasions per school year with standardized spelling tasks (in February and in June). After two school years, the children participated in four waves of spelling assessments, therefore. All spelling assessments were administered in classroom setting, and were carried out by the daily teacher of the children. This is in accordance with the manual of the standardized task and reliability was determined based on this procedure (De Wijs et al., 2010). The children received a short instruction before the assessment started. All items in the spelling assessments were scored manually by a trained test administrator (right or wrong; mostly by the daily teacher of the children).

There were three different spelling tasks (Single word dictation task, Sentence context dictation task, and Spelling decision task), which were conducted in different compositions based on measurement wave and on individual differences between the children (cf. Rasch, 1960). Each measurement wave consisted of two parts: During the first part, the global level of the individuals was determined by a dictation task which assesses random sample items throughout a continuous scale (single word dictation in Grade 1 and Sentence dictation in Grade 2). After determination of the spelling level, classes were divided into a group of poor spellers and a group of good spellers. The cut off score for the classification is determined by extensive accuracy tests of the standardized assessment method (De Wijs et al., 2010). Both groups got different, level adapted, more specific spelling tasks for the second part of the concurrent measurement wave. This differentiated testing is called multi stage testing. Test results between and within individuals can be compared on one and the same scale, regardless of the items that have been administered for individuals (see Rasch, 1960).

Halfway Grade 1, all children were first assessed in a global Single word dictation task. In the same measurement wave, all participating children conducted a second, more specific, Single word dictation task. Now, the poor spellers received other (level adapted) words in the dictation task than the good spellers. In total, all children wrote down 50 target words during this measurement moment.

By the end of Grade 1, again multi stage testing was conducted. All children were first assessed in a Sentence context dictation task. Right thereafter, poor and good spellers received different words in a second Sentence dictation task. In total, all children wrote down 50 target words during this measurement wave.

Halfway Grade 2, all children were first assessed in a Sentence context dictation task. Subsequently, the poor spellers got a second Sentence dictation task, while the good spellers received a Spelling decision task. In total, 50 target words were assessed for all children by the end of this measurement moment.

By the end of Grade 2, the Sentence dictation task followed by a second Sentence dictation task for poor spellers and a Spelling decision task for good spellers, similar to halfway Grade 2.

2.4. Analytic approach

We analyzed the data with LISREL longitudinal latent path modeling (Jöreskog & Sörbom, 1996) to examine both the word reading development and the spelling development during the first two years of elementary school. Autoregression models were used, since they allow the incorporation of other sources of variance, and therefore were found particularly suited to examine whether different determinants affect interindividual variance at different times (Bast & Reitsma, 1997). This made these models suitable for examining the longitudinal development of word reading and spelling, their integrated development (in a cross-lagged latent panel model), and their cognitive and linguistic precursors in the current study.

Cross-lagged latent panel models are particularly useful to examine predictive regression relations among latent constructs over time (Little, 2013). These models analyze individual differences expressed as change over time, and therefore fit very well with our research aims. In contrast to growth modeling studies, the current study focused more on between subject differences instead of within subjects changes like rate and shape of change. In the cross-lagged latent panel model, the residual covariances between the endogenous variables (word reading and spelling) were freely estimated. This means that we take into account that the variables at one time point might share some common `cause` not explained by the specified predictors. The latent variables of reading and spelling, and the latent variable of rapid naming consisted of one indicator each, namely the observed variable of the skill at the specific time point. The latent variable of early literacy consisted of three indicators, namely grapheme-to-phoneme knowledge, initial phoneme isolation, and word segmentation. The error terms of each variable were set equal to 0. The factor loadings for each latent variable were fixed to 1, in order to identify and estimate the model. For the latent variable of early literacy, the grapheme-to-phoneme skill was fixed to 1, while the other two skills were freely estimated. Time was used as a fixed factor in the current design, meaning that variables later in time were considered not to influence variables earlier in time.

In the conceptual model, a measurement moment of one construct was considered to be influenced by the directly preceding measurement of the same construct (autoregression). The child characteristics measured in kindergarten were added to the models in order to measure the predictive value of the child characteristics in kindergarten on the later literacy development. Both spelling and word reading development were first considered independently in separated models, and subsequently the reciprocal relationships of the cross-lags were analyzed in an integrated model. We deleted, one-by-one, the non-significant prediction paths starting with the path that showed the weakest relation (the model was re-ran after every modification). Only those influences significant at $\alpha < 0.05$ were presumed in the models. This is in line with Little (2013) to delete non-significant effects which do not contribute to the prediction of the construct that was tested in the model, unless strong theoretical expectations were formulated on specific paths. After testing the hypothesized models, the Modification index and the associated Expected parameter changes were consulted in LISREL for plausible modifications (see Little, 2013; Saris, Satorra, & Van der Veld, 2009). Modifications were accepted if the expected change was both significantly contributing to the fit of the model (MI ≥ 3.84, which refers to significance improvement of the chi^2 at $\alpha = 0.05$), and was theoretically plausible. The fit of the models was evaluated using a chi-square (χ^2) test. Because of the large sample size in the current study, the power to reject the model was too high to only use χ^2 statistic as a decisive criterion (Jaccard & Wan, 1996). Therefore, in addition the Root Mean Square Error of Approximation (RMSEA) and the relative chi-square (χ^2_{rel}), calculated as the ratio of the chi-square with the degrees of freedom, were taken into account. The RMSEA adds an additional correction factor for the effect of sample size and the degrees of freedom. In addition, we calculated the confidence interval around the point estimate of the RMSEA, to provide some more information about the range in which the true value may fall (Little, 2013). The critical value for RMSEA was set on < 0.06 to be considered good fit (Tabachnick & Fidell, 2014), on 0.06–0.08 to provide acceptable model fit, and 0.08–0.10 for mediocre fit (Little, 2013). The relative chi-square should be lower than 3 to be considered good fit (Carmines & McIver, 1981).

Table 1
Means and standard deviations of cognitive and linguistic precursor measures in kindergarten (N = 487).

Measure	Mean (*SD*)	Max. score	Skewness (SE)	Kurtosis (SE)
Grapheme-phoneme knowledge	18.34 (7.45)	34	− 0.14 (0.11)	− 0.72 (0.22)
Initial phoneme isolation	8.35 (1.83)	10	− 1.84 (0.11)	4.17 (0.22)
Word segmentation	4.35 (2.55)	10	0.01 (0.11)	− 0.57 (0.22)
Rapid naming	40.71 (8.90)	66	− 0.01 (0.11)	0.22 (0.22)
Visual short term memory	8.19 (3.02)	15	− 0.40 (0.11)	− 0.28 (0.22)
Verbal short term memory	15.34 (3.12)	20	− 0.85 (0.11)	0.50 (0.22)
Active vocabulary	14.03 (4.42)	25	− 0.45 (0.11)	0.13 (0.22)

3. Results

Prior to analysis, the data were examined for missing values. None of the included variables missed > 2.5% of the values, and the missing pattern of the current dataset was considered at random (Little's MCAR test: $\chi^2(83) = 103.95, p = 0.06$). The dataset was perfectly suitable for classic analyses and for LISREL path modeling. The parameters were estimated using the Full Information Maximum Likelihood (FIML) approach in LISREL (Enders, 2010; Little, 2013). The means, standard deviations, and skewness and kurtosis values of the kindergarten measures are presented in Table 1. Skewness and kurtosis were within the range of determining normality (skewness reference value > 2.1; kurtosis reference value > 7.1; Kim, 2013; West, Finch, & Curran, 1995). The four measurements of word reading and the four measurements of spelling are presented in Table 2. Paired sample *t*-tests showed that the mean scores steadily increase between every consecutive measurement moment for both word reading and spelling from halfway Grade 1 through the end of Grade 2 (all p's < 0.001).

The moderate correlations of the kindergarten measures in Table 3 confirmed that all kindergarten measures were related, nonetheless measured independent skills. Grapheme-phoneme knowledge, initial phoneme isolation, and word segmentation were most related, as expected, for all measuring components of early literacy. As is additionally shown by Table 3, all kindergarten measures were related to the outcome measures of both word reading and spelling. Active vocabulary was not correlated with word reading and the small correlation with spelling disappeared by the end of Grade 2. Word reading and spelling were correlated measures.

3.1. Prediction of word reading development

To address the first research question about the stability of word reading development and the determination by kindergarten measures, first a longitudinal developmental model of the repeated measures of word reading was built. Paths were specified between the subsequent measures of word reading. Second, the predictive values of the kindergarten measures on the latent autoregression model were analyzed. We first specified paths considering prior variables to have direct causal influences on variables at the immediately following test time. Therefore, the kindergarten precursors were released on the first measurement moment, halfway Grade 1. If the model asked for further independent contributions, we added the additional predictive paths (Modification indices). However, no plausible additional paths were proposed by the model. Only lags forward in time were considered. It was assumed that influences of the kindergarten abilities would be transmitted through the developmental path. This resulted in the model depicted in Fig. 2a. The model shows high autoregression in word reading development through Grade 1 and 2 (see Table 4 for the estimated parameter effects). The path model was predicted by early literacy and rapid naming. Although informative, this model had a mediocre fit to the data, $\chi^2(14) = 68.85, p < 0.001$, RMSEA = 0.090, RMSEA 90% CI = 0.069–0.110, $\chi^2_{rel} = 4.92$ (Little, 2013).

3.2. Prediction of spelling development

To address the research question about the stability of spelling development and the determination by kindergarten measures, first, the latent autoregressive model was also evaluated for the spelling measures. Second, the kindergarten measures were added to the autoregressive model, starting with early literacy. After the contribution of early literacy, no further kindergarten measures were contributing to the prediction model. Third, modification indices in LISREL suggested a direct path from early literacy to the spelling measurement by the end of Grade 2. This suggestion was assumed plausible within the context of

Table 2
Means and standard deviations of word reading and spelling (N = 487).

Measurement moment	Word reading (correct per minute)			Spelling (standardized)		
	Mean (*SD*)	Skewness (SE)	Kurtosis (SE)	Mean (*SD*)	Skewness (SE)	Kurtosis (SE)
Halfway Grade 1	51.01 (28.99)	1.30 (0.11)	2.57 (0.22)	111.02 (6.67)	0.23 (0.11)	− 0.12 (0.22)
Card 1	31.30 (18.29)					
Card 2	22.47 (17.25)					
End Grade 1	114.11 (53.62)	0.56 (0.11)	− 0.33 (0.22)	116.01 (5.87)	0.56 (0.11)	0.52 (0.22)
Card 1	50.41 (19.62)					
Card 2	36.94 (18.98)					
Card 3	23.59 (13.72)					
Halfway Grade 2	168.37 (60.71)	0.05 (0.11)	− 0.62 (0.22)	122.20 (6.81)	0.32 (0.11)	0.05 (0.22)
Card 1	69.71 (20.67)					
Card 2	57.62 (22.89)					
Card 3	40.86 (18.80)					
End Grade 2	191.71 (59.54)	− 0.06 (0.11)	− 0.40 (0.22)	123.51 (7.63)	0.80 (0.11)	1.39 (0.22)
Card 1	75.73 (19.13)					
Card 2	64.82 (22.19)					
Card 3	49.30 (18.66)					

Note. For some participants, the schools did provide the sum scores of the Word Reading cards, but not the separate scores of the cards. This explains the small discrepancies between means of card and means of sum scores.

Table 3
Correlations between precursor measures, word reading efficiency, and spelling ability.

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. GPK	–														
2. IPI	0.493**	–													
3. WS	0.519**	0.501**	–												
4. RAN	0.394**	0.263**	0.258**	–											
5.STMvis	0.234**	0.130**	0.219**	0.187**	–										
6.STMverb	0.254**	0.386**	0.288**	0.203**	0.127**	–									
7. VOC	0.173**	0.202**	0.225**	0.257**	0.076	0.259**	–								
8. WRM1	0.539**	0.278**	0.328**	0.380**	0.229**	0.239**	0.090*	–							
9. WRE1	0.462**	0.242**	0.261**	0.357**	0.222**	0.220**	0.049	0.846**	–						
10.WRM2	0.389**	0.178**	0.177**	0.362**	0.185**	0.189**	0.003	0.721**	0.883**	–					
11.WRE2	0.347**	0.142**	0.132**	0.357**	0.180**	0.178**	− 0.005	0.664**	0.845**	0.939**	–				
12. SPM1	0.441**	0.316**	0.322**	0.231**	0.155**	0.274**	0.196**	0.462**	0.401**	0.317**	0.289**	–			
13. SPE1	0.408**	0.284**	0.306**	0.245**	0.185**	0.268**	0.263**	0.506**	0.502**	0.419**	0.393**	0.582**	–		
14. SPM2	0.486**	0.268**	0.284**	0.273**	0.207**	0.238**	0.110*	0.578**	0.623**	0.581**	0.563**	0.583**	0.617**	–	
15. SPE2	0.500**	0.230**	0.261**	0.303**	0.248**	0.177**	0.046	0.614**	0.652**	0.623**	0.611**	0.426**	0.472**	0.746**	–

Note. *$p < 0.05$; **$p < 0.01$; GPK = grapheme-phoneme knowledge; IPI = initial phoneme isolation; WS = word segmentation; RAN = rapid naming; STMvis = visual short term memory; STMverb = verbal short term memory; VOC = vocabulary; WRM = Word Reading Middle of the year; WRE = Word Reading End of the year; SPM = Spelling Middle of the year; SPE = Spelling End of the year.

our conceptual model, and it significantly increased the fit of the model to the data. Fig. 2b shows the resulting model. High autoregression in the development of spelling ability has been shown, and the developmental path is predicted by early literacy (also see Table 4). Although the model visualized the spelling development, the model had a mediocre fit to the data, χ^2 (9) = 42.13, $p < 0.001$, RMSEA = 0.087, RMSEA 90% CI = 0.062–0.110, $\chi^2_{rel} = 4.68$ (Little, 2013). The model was considered sufficient for exploration of the predictive and developmental relationships, but results should be interpreted with caution, therefore.

3.3. Integration of word reading development and spelling development

The main focus of the study was the integrated development of word reading and spelling. A combined path model was built with both word reading efficiency and spelling, as measures of literacy development across the first two grades of elementary school. Our conceptual model of the integrated development was based on the combination of the singular basic simplex change processes, our theoretical expectations, and on the guidance of previous work.

First, paths were specified between the subsequent measures of the same construct. Second, the variables at concurrent measurement moments were allowed to covary. The residual covariances between the endogenous variables (word reading and spelling) were freely estimated, and therefore, allowed to associate. This association was not significant at the end of Grade 2 and, therefore, was not retained in the model. Third, the first order cross-lag relationships between word reading and spelling were analyzed over and above the autoregressive relations (cross-lag panel model). That means, we analyzed the effects of word reading at time t on spelling at time $t + 1$ and vice versa,

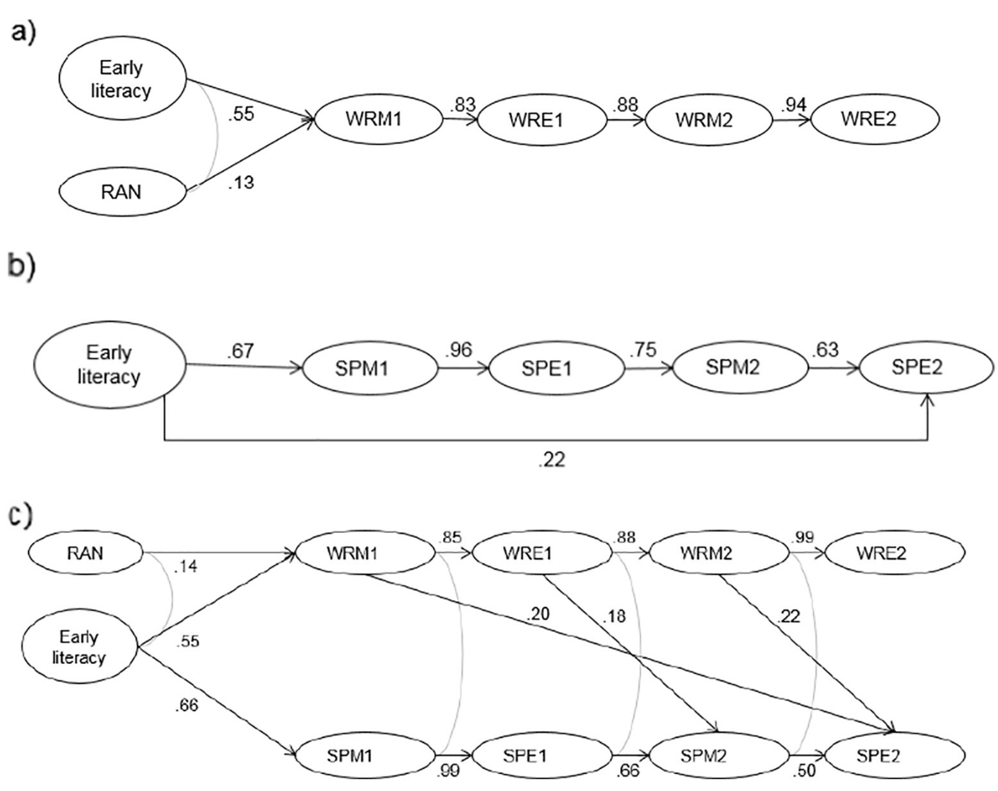

Fig. 2. Latent autoregressive prediction models of a) word reading development, b) spelling development, and c) the developmental interrelationships between word reading development and spelling development in a first order cross-lags latent path model. The completely standardized path coefficients are presented in the models. All developmental models have been analyzed in combination with kindergarten precursors. The residual variances among the measurements at one measurement moment were allowed to associate and the early literacy and RAN were allowed to correlate, as indicated by the lines in grey. SP = Spelling; WR = Word Reading; RAN = Rapid Naming; M = Mid-year; E = End of year.

Table 4
Estimated parameter effects for the singular model of both spelling and word reading, and for the integrated model.

Parameters	Singular models		Integrated model
	Reading	Spelling	
β SPM1 - SPE1		0.85 (0.07)	0.89 (0.07)
β SPE1 - SPM2		1.19 (0.08)	1.03 (0.09)
β SPM2 - SPE2		0.64 (0.08)	0.53 (0.08)
β WRM1 - WRE1	1.51 (0.07)		1.56 (0.07)
β WRE1 - WRM2	0.97 (0.03)		0.97 (0.03)
β WRM2 - WRE2	0.95 (0.02)		0.95 (0.02)
ψ SPM1			14.48 (1.92)
ψ SPM2			0.25 (1.33)
ψ SPM2			19.48 (4.15)
ψ SPE2			19.57 (1.80)
ψ WRM1			525.94 (48.02)
ψ WRE1			831.09 (97.55)
ψWRM2			819.84 (73.81)
ψ WRE2			385.09 (73.80)
ψ WRM1 - SPE2			34.30 (6.26)
ψ WRE1 - SPM2			22.70 (5.42)
ψWRM2 - SPE2			26.43 (6.18)
α Early literacy – SPM1		3.03 (0.33)	0.52 (0.05)
α Early literacy – SPE2		1.43 (0.37)	
α Early literacy – WRM1	10.15 (1.08)		2.59 (0.25)
α RAN – WRM1	0.44 (0.15)		0.47 (0.14)
ϕ Early literacy - RAN	6.49 (0.89)		25.20 (3.10)
R^2 SPM1		0.45	0.43
R^2 SPE1		0.92	0.99
R^2 SPM2		0.57	0.61
R^2 SPE2		0.58	0.65
R^2 WRM1	0.38		0.39
R^2 WRE1	0.70		0.72
R^2 WRM2	0.77		0.77
R^2 WRE2	0.88		0.89

Note. Standard errors are in parentheses. β = unstandardized β values are presented in this table; Ψ = residuals of the latent variables and the residual covariances between two concurrent variables; ϕ = the covariance between the two observed variables which remained in the model; R^2 = Proportion variance explained; WRM = Word Reading Middle of the year; WRE = Word Reading End of the year; SPM = Spelling Middle of the year; SPE = Spelling End of the year.

starting at measurement moment 1. Only lags forward in time were considered. Fourth, the precursors measured in kindergarten were added to the integrated model. First, the independent contributions on the first longitudinal measurement moment were considered and added as additional paths. Thereafter, possible additional predicting paths to the other measurement moments were considered, using the modification indices. A direct path from the first word reading measurement (halfway Grade 1) to the last spelling measurement (end of Grade 2) was suggested. This suggestion was assumed plausible within the context of our conceptual model, and it significantly increased the fit of the model to the data. Although this path was not hypothesized, it was decided to be of theoretical plausible relevance. It was preserved in the resulting model, therefore. See Fig. 2c for the resulting model. Also see Table 4 for the estimated parameter effects for the integrated model. The model showed a clear pattern of cross-lag relationships from word reading to spelling. The reversed reciprocal effect was not found to contribute to the fit of the model. The cross-lag effects were relatively constant over time, which suggests that the predictive value of word reading on spelling is developmentally stable during the first two years of literacy development. It should be noted that the contribution of the cross-lag effects were of a smaller magnitude than the contribution of the autoregression of the spelling development. Word reading development was best predicted by early literacy, and to a lesser extent by rapid naming. Spelling development was predicted by early literacy. The model fitted the data adequately and was considered acceptable, χ^2 (38) = 98.85, $p < 0.001$, RMSEA = 0.057, RMSEA 90% CI = 0.044–0.071, $\chi^2_{rel} = 2.60$. The completely standardized path coefficients are presented.

4. Discussion

The present longitudinal study examined the singular and integrated word reading and spelling development during the first two grades of Dutch primary education in relation to kindergarten precursor measures. Results show that both spelling and word reading development were highly stable and consistently autoregressive in nature. Both spelling and word reading development were best predicted by kindergarten measures of early literacy (i.e., phonemic awareness, grapheme-to-phoneme knowledge). Word reading development was additionally predicted by rapid naming. The integrated model for word reading and spelling development showed that reading and spelling were related, and that word reading level predicted subsequent spelling level in Grade 2 over and above the autoregressive prediction of spelling itself. The current study in the transparent Dutch orthography adds to our knowledge on the integrated early reading and spelling development in orthographies other than the highly opaque English orthography, strengthening our knowledge about general underlying principles in literacy development.

The singular models show that, in line with the findings in the literature (e.g., Furnes & Samuelsson, 2011; Juul et al., 2014; Pinto et al., 2015; Steacy et al., 2014), word reading efficiency was predicted mostly by itself in an autoregressive model. Word reading development was predicted from kindergarten by early literacy (phonemic awareness, grapheme-to-phoneme knowledge) and rapid naming, in line with the literature (e.g., Furnes & Samuelsson, 2011; Moll et al., 2014).

Spelling development was also best predicted by itself in an autoregressive model, and by early literacy measured in kindergarten. These results suggest that knowing which phoneme corresponds to which grapheme, together with the ability to isolate and segment the phonemes in spoken words (i.e., phonological skills), is a prerequisite for both learning to read and spell. This finding is in line with previous predictive studies on either reading and spelling development in both transparent and opaque orthographies (e.g., Furnes & Samuelsson, 2009; Georgiou et al., 2012). That implies that both the grapheme-to-phoneme conversion (as required in reading) and the phoneme-to-grapheme conversion (as required in spelling) could be predicted from a strong early literacy level in kindergarten.

After the contribution of early literacy, no other kindergarten precursors were found to contribute to the prediction of spelling development in Grade 1 and 2. This finding is in line with previous studies showing that early literacy is a stronger predictor of early spelling development than rapid naming, at least in transparent orthographies (e.g., Furnes & Samuelsson, 2011; Landerl & Wimmer, 2008). However, it should be noted that Lervåg and Hulme (2010), did find an influence of rapid naming, and to a lesser extent of short term memory on the intercept (but not on the growth) of spelling development. Verhagen et al. (2010) also found a contribution of rapid naming on the prediction of spelling, and to a lesser extent of vocabulary, although it should be mentioned that they did not measure grapheme-phoneme knowledge as a precursor in their study. The contribution of our kindergarten rapid naming measure was relatively small as compared to the high contribution of rapid naming in some previous literature in transparent orthographies (e.g., Vaessen & Blomert, 2013). We should explicitly mention here that in the current study only one measure of nonalphanumeric stimuli was used. Nonalphanumeric stimuli are preferred in prereaders, since the stimuli should be 'highly familiar' to tap into the automated retrieval skills (e.g., Kirby et al., 2010; Lervåg & Hulme, 2009). It is arguable, however, that a rapid naming construct compiled out of more than one RAN measurement could have given higher contributions to the model.

Over and above the contribution of early literacy that was provided from kindergarten to halfway Grade 1, and indirectly through the longitudinal model, early literacy directly contributed to the prediction of spelling performances by the end of Grade 2. This suggests that phonemic awareness and grapheme-phoneme knowledge remain stable contributors of spelling development over the years. The cross-sectional concurrent study of Vaessen and Blomert (2013) also found stable contributions of cognitive determinants on spelling, while the impact of early literacy on the later reading development seemed to decrease.

The integrated model describes the developmental relations between word reading and spelling in the transparent Dutch orthography. Firstly, the integrated model of both word reading and spelling shows that word reading and spelling are related to each other. The model fitted to the data very well. Remarkably, however, if both reading and spelling development were modelled independently, the fit of the two singular models was acceptable but mediocre. Therefore, the integrated literacy model was considered a better representation of the data. Early development of word reading and spelling can be considered as integrated skills that reinforce each other's process of development.

Secondly, the autoregressive relations within the domains of word reading and spelling was stronger than the cross-domain contributions, demonstrating that reading and spelling are related but different processes. This finding is similar to findings in previous research in opaque orthographies (Abbott et al., 2010; Ahmed et al., 2014; Foorman et al., 2011; Shanahan, 2006). In addition, the autoregressive development of word reading was found to be more consistent as compared to the autoregression of spelling development (which decreases over time), which might be interpreted as a representation of the strong grapheme-to-phoneme connections in reading as compared to the weaker phoneme-to-grapheme connections in spelling (see Van Orden et al., 1990). This finding is in line with other studies on the consistency and the stability of literacy development (e.g., Pinto et al., 2015).

Thirdly, and more into detail of the cross-pathways in the integrated model of word reading and spelling development, word reading was longitudinally contributing to subsequent spelling skills. The current findings provide support for the notion in previous developmental models (see Frith, 1985) that reading and spelling were not only related to each other; children also apply the knowledge that they have learned in reading to their further spelling development. In Grade 2, the cross-domain predictive value of word reading on spelling was developmentally stable. No contribution was found from word reading halfway Grade 1 to the subsequent spelling measure by the end of Grade 1. This indicates that spelling was not predicted by word reading during the first year of formal instruction. Indeed, the high autoregression from spelling performances halfway Grade 1 to the end of Grade 1 shows that spelling largely grows from its own previous performances during this first year, with no additional contributions of word reading skills. In addition, spelling in Grade 1 is also highly predicted by the kindergarten measure of early literacy. This early literacy measure might be overlapping with the relevant components within the initial word reading measure (i.e., phonemic awareness and the strength of the grapheme-phoneme connections), and therefore takes account for the prediction of spelling in first grade. In addition to the lag from end Grade 1, word reading halfway Grade 1 also directly contributed to the prediction of spelling by the end of Grade 2, emphasizing the stability of the cross-domain influence of reading on spelling development. Furthermore, this extra contribution of early word reading to spelling more than a year later, adds to the theoretical suggestion by Frith (1985) that reading is the pacemaker for orthographic spelling. Word reading becomes predictive during later spelling development, because the words to be spelled become more complex. Therefore, higher levels of orthographic representations are required (see also Ellis, 1994).

We found no evidence for the reciprocal influence of spelling on subsequent word reading. This means that, whereas children apply their reading skills in order to improve their spelling skills, this developmental advantage was not found from spelling to reading. The current finding is in line with other studies on the integrated development of reading and spelling (Ahmed et al., 2014; Caravolas et al., 2001; Desimoni et al., 2012). However, there are also studies that did find a reciprocal relation (Abbott et al., 2010; Ellis & Cataldo, 1990; Lerkkanen et al., 2004; Pinto et al., 2015). We found four explanations for the divergent findings. Firstly, the contrasting findings could be partly explained by the different characteristics of orthographies, in terms of asymmetries in backward and forward conversions (see Bosman & Van Orden, 1997). Wimmer and Landerl (1997) suggested that orthographies with consistent grapheme-to-phoneme relations but less consistent phoneme-to-grapheme relations (like Dutch) have the advantage of reading experience while learning to spell. It could be argued that spelling development also supports the phonemic awareness and grapheme-phoneme connections (phonologic mediation is fundamental to both reading and spelling; Bosman & Van Orden, 1997), so the reciprocal advantage could be evenly present. However, the high consistency in the reading process might not need the additional support by spelling, while the more inconsistent spelling process takes advantage of the extra support by reading (also found by Shanahan, 2006).

Secondly, differences in findings can also be provoked by different ways of measuring reading and spelling constructs. Some studies included reading comprehension measures (e.g., Leppänen et al., 2006; Lerkkanen et al., 2004) or free writing exercises (e.g., Pinto et al., 2015) already during the very early literacy development, whereas in other studies, the absolute measure of word decoding and word spelling were conducted. The current study assembled to the absolute measures. In accordance with for example Babayiğit and Stainthorp (2010) and Pinto et al. (2015), timed word reading instead of word reading accuracy was analyzed in the longitudinal models.

A third explanation for the divergent findings is the contribution of the kindergarten precursor measures in the models. For example, Leppänen et al. (2006) found an early influence of kindergarten spelling measures to reading during the very beginning of literacy development which diminished during later development. They, however, showed that the early influence disappeared as soon as kindergarten phonological awareness was also taken into account. This suggested that early influence of spelling to reading was explained by phonological awareness skills. In the current model, phonological awareness skills were represented by the early literacy measure in kindergarten.

Lastly, the time window of interest is another explanation for different findings. As has been proposed by Frith (1985), reading and spelling might take turn in their influence to each other over developmental course. Spelling-to-reading benefits often have been found at the initial phase of learning the alphabetic principle (Leppänen et al., 2006; Shahar-Yames & Share, 2008). The children in our study mainly mastered the alphabetic principle within the first half a year of instruction. Therefore, they were past the initial code learning phase right before the first measurement moment. The focus was already on word reading efficiency instead of accuracy. The connections between graphemes and phonemes may have resulted into orthographic representations, which may have provided children with orthographic knowledge instead of discrete phonological decoding strategies. In more opaque orthographies like English, children spend longer on learning the alphabetic principle and on becoming accurate in reading. They may take more advantage of practicing spelling for further reading development, therefore.

Adjacent to that argument, in future research, it is valuable to study the early literacy development earlier and with more fine-grained time intervals. In the current study we have focused on development during formal early literacy instruction. We first assessed word decoding from the time that all letters had been taught (halfway Grade 1). Thereafter, three consecutive measurement moments followed each half-year period (end Grade 1, halfway Grade 2, end Grade 2). Because it is possible that children already have some word reading and spelling abilities before formal literacy instruction starts, taking kindergarten emergent skills of reading and (invented) spelling into consideration could have shed light on the interrelation during the earliest phases of development. The current study has some more limitations which might be addressed in future research. Firstly, some decisions could have limited the generalizability in terms of sample, method, and population. The current study may generalize only to Dutch children in the Netherlands, and we restricted to one type of reading and of spelling assessment, and all schools in the current study made use of the same, phonics based, Dutch literacy curriculum. Although an advantage of this design is that it contributes to the control and stability of the learning environment in all participants, additional research would be needed to establish whether the same relations hold in other orthographies and in other reading curricula (e.g., McGeown, Johnston, & Medford, 2012). Furthermore, we did not report word reading accuracy separately from word reading speed in the current study. Instead, in our reading efficiency measure, both speed and accuracy is inherently taken into account. Although this efficiency measure of word reading seems most meaningful in the Dutch transparent orthography (Verhoeven & Van Leeuwe, 2009), it makes it more difficult to compare the outcomes of the current study to other studies on word reading development. Thirdly, it would be relevant to study other precursor measures like morphological awareness and verbal processing skills, and more specific, the detailed phonological input processing skills. Speech perception abilities have been assumed to be related to the development of phonological awareness skills, which, in turn, are related to both reading and spelling (e.g., Carroll et al., 2003; Janssen, Segers, McQueen, & Verhoeven, 2016). Lastly, for future research, individual differences in handwriting fluency and writing quality should be considered in the longitudinal model to learn more about their influences in literacy development (see e.g., Ahmed et al., 2014).

Some practical implications add to the theoretical implications of the current study. Firstly, kindergarten precursors of learning to read are also relevant precursors of learning to spell. These early markers of individual differences in both word reading and spelling can help teachers to optimally adapt to the children's needs. Instead of waiting them to fail, Grade 1 teachers can immediately provide differentiated instruction to children at risk for later literacy problems. Secondly, our results show high stability of individual differences in both word reading and spelling development, which remain stable and more important than the cross-domain relations in the integrated model of word reading and spelling. These results confirm that both specific instruction and practising for spelling and specific instruction and practising for reading are important in primary education for balanced support of the total literacy development. Thirdly, our results show that word reading and spelling are related and that children apply their word reading knowledge to support their spelling development. The grapheme-to-phoneme connections are not only found to be more consistent, this direction is also more heavily trained than the phoneme-to-grapheme connections. The latter is obvious, because reading is done more than writing. However, additionally, the emphasis of early literacy instruction, at least in phonics based instruction in the Netherlands, is more on reading than on spelling. This one-sided emphasis enhances the asymmetry between spelling and reading development. Once in place, the asymmetry is suggested to be self-perpetuating (Bosman & Van Orden, 1997). To strengthen the spelling development, emphasis should be on building stronger phoneme-to-grapheme connections. Both domains of literacy instruction (i.e., both reading and spelling instruction) should be integrated and matched to each other to let children benefit from the newly learned knowledge across domains and to trigger and strengthen the bidirectional connections between phonemic, graphemic, and semantic knowledge.

In conclusion, the current study showed that word reading, spelling, and their determinants are closely linked to each other, already during the early phases of literacy development. Our results show that word reading and spelling are related, and that word reading is supportive for the prediction of subsequent spelling development in Grade 2. Not considering writing development in literacy research and education is at the expense of complete understanding of literacy development.

Acknowledgements

We thank dr. Ben J. Pelzer for his helpful advice about the statistical models.

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. http://dx.doi.org/10.1037/a0019318.

Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology, 106*(2), 419–434. http://dx.doi.org/10.1037/a0035692.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23*(5), 300–316. http://dx.doi.org/10.1177/07419325020230050501.

Babayiğit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing, 23*(5), 539–568. http://dx.doi.org/10.1007/s11145-009-9173-y.

Bast, J., & Reitsma, P. (1997). Mathew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32*(2), 135–167. http://dx.doi.org/10.1207/s15327906mbr3202_3.

Bekebrede, J., Van der Leij, A., & Share, D. L. (2009). Dutch dyslexic adolescents: Phonological-core variable-orthographic differences. *Reading and Writing, 22*(2), 133–165. http://dx.doi.org/10.1007/s11145-007-9105-7.

Bosman, A. M. T., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.). *Learning to spell: Research, theory, and practice across languages* (pp. 173–194). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bosse, M. L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: A cross-sectional study. *Journal of Research in Reading, 32*(2), 230–253. http://dx.doi.org/10.1111/j.1467-9817.2008.01387.x.

Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist, 9*(1), 3–14.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45*, 751–774. http://dx.doi.org/10.1006/jmla.2000.2785.

Caravolas, M., Lervåg, A., Defior, S., Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*(8), 1398–1407. http://dx.doi.org/10.1177/0956797612473122.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavsky, M., Onochie-Quintanilla, E., ... Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*(6), 678–686. http://dx.doi.org/10.1177/0956797611434536.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. W. Bohrnstedt, & E. F. Borgatta (Eds.). *Social measurement* (pp. 65–116). Beverly Hills, CA: Sage.

Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology, 39*(5), 913. http://dx.doi.org/10.1037/0012-1649.39.5.913.

Centraal Bureau voor de Statistiek [Statistics Netherlands] (2013, December 12). Retrieved from: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71822NED&D1=0-1&D2=0&D3=0&D4=04&D5=a&D6=0&D7=l&HDR=T,G2,G1,G6,G5,G4&STB=G3&VW=T.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204. http://dx.doi.org/10.1037//0033-295x.108.1.204.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733 (doi: 0022-O663/9O/).

De Graaff, S., Hasselman, F., Verhoeven, L., & Bosman, A. M. (2011). Phonemic awareness in Dutch kindergartners: Effects of task, phoneme position, and phoneme class. *Learning and Instruction, 21*(1), 163–173. http://dx.doi.org/10.1016/j.learninstruc.2010.02.001.

De Jong, P. F., & Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*(3), 450. http://dx.doi.org/10.1037/0022-0663.91.3.450.

De Wijs, A., Kamphuis, F., Kleintjes, F., & Tomesen, M. (2010). *Wetenschappelijke verantwoording spelling voor groep 3 tot en met 6 [Scientific justification spelling test of Grade 1 to Grade 4]*. Arnhem, Netherlands: Stichting Cito Instituut voor Toetsontwikkeling.

Desimoni, M., Scalisi, T. G., & Orsolini, M. (2012). Predictive and concurrent relations between literacy skills in Grades 1 and 3: A longitudinal study of Italian children. *Learning and Instruction, 22*(5), 340–353. http://dx.doi.org/10.1016/j.learninstruc.2012.02.002.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22*(6), 356–365. http://dx.doi.org/10.1177/002221948902200606.

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*(3), 19–36. http://dx.doi.org/10.1097/00011363-200020030-00005.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188. http://dx.doi.org/10.1207/s1532799xssr0902.

Ellis, N. (1994). Longitudinal studies of spelling development. In G. D. A. Brown, & N. C. Ellis (Eds.). *Handbook of spelling: Theory, process, and intervention* (pp. 3–34). Chichester: John Wiley & Sons Ltd.

Ellis, N., & Cataldo, S. (1990). The role of spelling in learning to read. *Language and Education, 4*(1), 1–28. http://dx.doi.org/10.1080/09500789009541270.

Enders, C. K. (2010). *Applied missing data analysis.* Guilford Publications.

Fletcher-Flinn, C. M., Shankweile, D., & Frost, S. J. (2004). Coordination of reading and spelling in early literacy development: An examination of the discrepancy hypothesis. *Reading and Writing, 17*(6), 617–644. http://dx.doi.org/10.1023/B:READ.0000044297.85675.f5.

Foorman, B. R., Arndt, E. J., & Crawford, E. C. (2011). Important constructs in literacy learning across disciplines. *Topics in Language Disorders, 31*, 73–83. http://dx.doi.org/10.1097/TLD.0b013e31820a0b86.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart, & J. Marshall (Eds.). *Surface dyslexia*. London: Erlbaum.

Furnes, B., & Samuelsson, S. (2009). Preschool cognitive and language skills predicting Kindergarten and Grade 1 reading and spelling: A cross-linguistic comparison. *Journal of Research in Reading, 32*(3), 275–292. http://dx.doi.org/10.1111/j.1467-9817.2009.01393.x.

Furnes, B., & Samuelsson, S. (2010). Predicting reading and spelling difficulties in transparent and opaque orthographies: A comparison between Scandinavian and US/Australian children. *Dyslexia, 16*(2), 119–142. http://dx.doi.org/10.1002/dys.401.

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21*(1), 85–95. http://dx.doi.org/10.1016/j.lindif.2010.10.005.

Gebauer, D., Enzinger, C., Kronbichler, M., Schurz, M., Reishofer, G., Koschutnig, K., ... Fink, A. (2012). Distinct patterns of brain function in children with isolated spelling impairment: New insights. *Neuropsychologia, 50*(7), 1353–1361. http://dx.doi.org/10.1016/j.neuropsychologia.2012.02.020.

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing, 25*(2), 321–346. http://dx.doi.org/10.1007/s11145-010-9271-x.

Geudens, A., & Sandra, D. (2003). Beyond implicit phonological knowledge: No support for an onset–rime structure in children's explicit phonological awareness. *Journal of Memory and Language, 49*(2), 157–182. http://dx.doi.org/10.1016/S0749-596X(03)00036-6.

Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9*(4), 351–365. http://dx.doi.org/10.1207/s1532799xssr0904_2.

Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models.* London, England: Sage Publications Ltd.

Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-114).* Thousand Oaks, CA: Sage Publications.

Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2016). *Transfer from implicit to explicit phonological abilities in first and second language learners.* Language and Cognition: Bilingualism1–18. http://dx.doi.org/10.1017/S1366728916000523.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 user's reference guide* (2nd ed.). Chicago, IL: Scientific Software International.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437–447. http://dx.doi.org/10.1037//0022-0663.80.4.437.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*(4), 243. http://dx.doi.org/10.1037//0022-0663.78.4.243.

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology, 106*(4), 1096. http://dx.doi.org/10.1037/a0037100.

Keuning, J., & Verhoeven, L. (2008). Spelling development throughout the elementary grades: The Dutch case. *Learning and Individual Differences, 18*(4), 459–470. http://dx.doi.org/10.1016/j.lindif.2007.12.001.

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution using skewness and kurtosis. *Restorative Dentistry & Endodontocs, 38*(1), 52–54. http://dx.doi.org/10.5395/rde.2013.38.1.52.

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology/Psychologie Canadienne, 49*(2), 103–110. http://dx.doi.org/10.1037/0708-5591.49.2.103.

Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly, 45*(3), 341–362. http://dx.doi.org/10.1598/RRQ.45.3.4.

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *Wetenschappelijke verantwoording DMT en AVI [Scientific justification three minutes test and analysis of individualization formats].* Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., ... Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*(6), 686–694. http://dx.doi.org/10.1111/jcpp.12029.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*(1), 150. http://dx.doi.org/10.1037/0022-0663.100.1.150.

Leppänen, U., Nieme, P., Aunola, K., & Nurmi, J. E. (2006). Development of reading and spelling Finnish from preschool to Grade 1 and Grade 2. *Scientific Studies of Reading, 10*(1), 3–30. http://dx.doi.org/10.1207/s1532799xssr1001_2.

Lerkkanen, M., Rasku-Puttonen, H., Aunola, K., & Nurmi, J. (2004). The developmental dynamics of literacy skills during the first grade. *Educational Psychology, 24*, 793–810. http://dx.doi.org/10.1080/0144341042000271782.

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20*(8), 1040–1048. http://dx.doi.org/10.1111/j.1467-9280.2009.02405.x.

Lervåg, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? *Scientific Studies of Reading, 14*(6), 485–513. http://dx.doi.org/10.1080/10888431003623488.

Little, T. D. (2013). *Methodology in the social sciences: Longitudinal structural equation modeling.* New York, USA: The Guilford Press.

McGeown, S. P., Johnston, R. S., & Medford, E. (2012). Reading instruction affects the cognitive skills supporting early reading development. *Learning and Individual Differences, 22*(3), 360–364. http://dx.doi.org/10.1016/j.lindif.2012.01.012.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*(2), 322–352. http://dx.doi.org/10.1037/a0026744.

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading, 13*(5), 359–382. http://dx.doi.org/10.1080/10888430903162878.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., ... Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction, 29*, 65–77. http://dx.doi.org/10.1016/j.learninstruc.2013.09.003.

Mommers, C. L., Verhoeven, L., Koekebacker, E., Linden Van der, S., Stegeman, W., & Warnaar, J. (2003). *Veilig leren lezen [Learning to read safely].* Tilburg, Netherlands: Zwijsen.

Morken, F., & Helland, T. (2013). Writing in dyslexia: Product and process. *Dyslexia, 19*(3), 131–148. http://dx.doi.org/10.1002/dys.1455.

Muter, V. (1998). Phonological awareness: Its nature and its influence over early literacy development. In C. R. Hulme, & R. Joshi (Eds.). *Reading and spelling: Development and disorders* (pp. 113–126). Hillsdale, NJ: Lawrence Erlbaum Associates.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide, 5.* Los Angeles: Muthén & Muthén.

Nelson, R., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice, 18*(4), 255–267. http://dx.doi.org/10.1111/1540-5826.00080.

Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94*(1), 1–17. http://dx.doi.org/10.1016/j.jecp.2005.11.006.

Nunn, A. (1998). *Dutch orthography.* Utrecht, the Netherlands: Center for Language Studies.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.). *Reading acquisition* (pp. 145–174). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Perfetti, C. A. (1997). The psycholinguistics of spelling and reading. In C. A. Perfetti, & L. Rieben (Eds.). *Learning to spell: Research, theory and practice across languages*. Mahwah, NJ: Lawrence Erlbaum Associates.

Pinto, G., Bigozzi, L., Tarchi, C., Gamannossi, B. A., & Canneti, L. (2015). Cross-lag analysis of longitudinal associations between primary school students' writing and reading skills. *Reading and Writing, 28*(8), 1233–1255. http://dx.doi.org/10.1007/s11145-015-9569-9.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56. http://dx.doi.org/10.1037//0033-295x.103.1.56.

Pugh, K., Sandak, R., Frost, S. J., Moore, D. L., Mencl, W. E., Dickinson, D., & Neuman, S. (2006). What does reading have to tell us about writing? Preliminary question and methodological challenges in examining the neurobiological foundations of writing and writing disabilities. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of writing research* (pp. 433–448). New York, NY: Guilford Press.

Rasch, G. (1960). *Probabilistic models for some intelligence and achievement tests.* Copenhagen: Danish Institute for Educational Research.

Saris, W. E., Satorra, A., & Van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling, 16*(4), 561–582. http://dx.doi.org/10.1080/10705510903203433.

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Word decoding development in incremental phonics instruction in a transparent orthography. *Reading and Writing*, 1–22. http://dx.doi.org/10.1007/s11145-017-9735-3.

Schaerlaekens, A., Kohnstamm, D., Lejaegere, M., De Vries, A., Peeters, L., & Zink, I. (1999). *Streefwoordenslijst voor 6-jarigen [Target vocabulary list for 6-year old children].* Lisse: Swets & Zeitlinger.

Schreuder, R., & Van Bon, W. H. (1989). Phonemic analysis: Effects of word properties. *Journal of Research in Reading, 12*(1), 59–78.

Shahar-Yames, D., & Share, D. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading, 31*(1), 22–39. http://dx.doi.org/10.1111/j.1467-9817.2007.00359.x.

Shanahan, T. (1984). Nature of the reading-writing relation: An exploratory multivariate analysis. *Journal of Educational Psychology, 76*(3), 466–477. http://dx.doi.org/10.1037//0022-0663.76.3.466.

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of writing research* (pp. 171–183). New York, NY: Guilford Press.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95–129. http://dx.doi.org/10.1006/jecp.1998.2481.

Steacy, L. M., Kirby, J. R., Parrila, R., & Compton, D. L. (2014). Classification of double deficit groups across time: An analysis of group stability from kindergarten to second grade. *Scientific Studies of Reading, 18*(4), 255–273. http://dx.doi.org/10.1080/10888438.2013.873936.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36*(3), 337–359.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics.* Pearson Education Limited.

Torppa, M., Georgiou, G. K., Niemi, P., Lerkkanen, M. K., & Poikkeus, A. M. (2016). The precursors of double dissociation between reading and spelling in a transparent orthography. *Annals of Dyslexia,* 1–21. http://dx.doi.org/10.1007/s11881-016-0131-5.

Treiman, R., Cassar, M., & Zukowski, A. (1994). What types of linguistic information do children use in spelling? The case of flaps. *Child Development, 65*, 1310–1329. http://dx.doi.org/10.1111/j.1467-8624.1994.tb00819.x.

Tucker, R., Castles, A., Laroche, A., & Deacon, S. H. (2016). The nature of orthographic learning in self-teaching: Testing the extent of transfer. *Journal of Experimental Child Psychology, 145*, 79–94. http://dx.doi.org/10.1016/j.jecp.2015.12.007.

Vaessen, A., & Blomert, L. (2013). The cognitive linkage and divergence of spelling and reading development. *Scientific Studies of Reading, 17*(2), 89–107. http://dx.doi.org/10.1080/10888438.2011.614665.

Van den Boer, M., De Jong, P. F., & Haentjens-van Meeteren, M. M. (2013). Modeling the length effect: Specifying the relation with visual and phonological correlates of reading. *Scientific Studies of Reading, 17*(4), 243–256. http://dx.doi.org/10.1080/10888438.2012.683222.

Van Orden, G. C., Bosman, A. M., Goldinger, S. D., & Farrar, W. T. (1997). A recurrent-network account of reading, spelling, and dyslexia. *Advances in Psychology, 121*, 522–538. http://dx.doi.org/10.1016/s0166-4115(97)80113-6.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97*(4), 488. http://dx.doi.org/10.1037/0033-295X.97.4.488.

Verhagen, W. G. M., Aarnoutse, C. A. J., & Van Leeuwe, J. F. J. (2010). Spelling and word recognition in Grades 1 and 2: Relations to phonological awareness and naming speed in Dutch children. *Applied PsychoLinguistics, 31*(01), 59–80. http://dx.doi.org/10.1017/S0142716409990166.

Verhelst, N. D., & Glas, C. A. W. (1995). The one-parameter logistic model. In G. H. Fischer, & I. W. Moelenaar (Eds.). *Rasch models. Foundations, recent developments, and applications*. New York: Springer-Verlag.

Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1995). OPLM: One-parameter logistic model. *Computer program and manual*. Arnhem, Netherlands: Stichting Cito toetsontwikkeling.

Verhoeven, L., & Van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading, 13*(3), 205–223. http://dx.doi.org/10.1080/10888430902851356.

Verhoeven, L., & Vermeer, A. (1986). *TAK: Taaltoets Allochtone Kinderen [Language test ethnic minority children].* Tilburg, Netherlands: Uitgeverij Zwijsen.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.). *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Newbery Park, CA: Sage.

Wimmer, H., & Landerl, K. (1997). How learning to spell German differs from learning to spell English. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.). *Learning to spell: Research, theory, and practice across languages* (pp. 81–96). New Jersey, United States: Laurence Erlbaum Associates.

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology, 94*(2), 272. http://dx.doi.org/10.1037//0022-0663.94.2.272.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3. http://dx.doi.org/10.1037/0033-2909.131.1.3.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science, 9*(5), 429–436. http://dx.doi.org/10.1111/j.1467-7687.2006.00509.x.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers, 28*(4), 504–515.